

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(i) APPLICANT: Bisgaard-Frantzen, Henrik  
Svendsen, Allan  
Borchert, Torben Vedel

(ii) TITLE OF INVENTION: AMYLASE VARIANTS

(iii) NUMBER OF SEQUENCES: 32

## (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Novo Nordisk of North America, Inc.  
(B) STREET: 405 Lexington Avenue, Suite 6400  
(C) CITY: New York  
(D) STATE: New York  
(E) COUNTRY: U.S.A.  
(F) ZIP: 10174-6401

## (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

## (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 08/600,656  
(B) FILING DATE: 13-FEB-1996  
(C) CLASSIFICATION:

## (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Lambiris, Elias J.  
(B) REGISTRATION NUMBER: 33,728  
(C) REFERENCE/DOCKET NUMBER: 4318.204-US

## (ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 212 867 0123  
(B) TELEFAX: 212 867 0298

## (2) INFORMATION FOR SEQ ID NO: 1:

## (1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 485 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

His His Asn Gly Thr Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Tyr  
1 5 10 15

Leu Pro Asn Asp Gly Asn His Trp Asn Arg Leu Arg Asp Asp Ala Ala  
20 25 30

Asn Leu Lys Ser Lys Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Trp  
35 40 45

Lys Gly Thr Ser Gln Asn Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr  
 50 55 60  
 Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly  
 65 70 75 80  
 Thr Arg Asn Gln Leu Gln Ala Ala Val Thr Ser Leu Lys Asn Asn Gly  
 85 90 95  
 Ile Gln Val Tyr Gly Asp Val Val Met Asn His Lys Gly Gly Ala Asp  
 100 105 110  
 Gly Thr Glu Ile Val Asn Ala Val Glu Val Asn Arg Ser Asn Arg Asn  
 115 120 125  
 Gln Glu Thr Ser Gly Glu Tyr Ala Ile Glu Ala Trp Thr Lys Phe Asp  
 130 135 140  
 Phe Pro Gly Arg Gly Asn Asn His Ser Ser Phe Lys Trp Arg Trp Tyr  
 145 150 155 160  
 His Phe Asp Gly Thr Asp Trp Asp Gln Ser Arg Gln Leu Gln Asn Lys  
 165 170 175  
 Ile Tyr Lys Phe Arg Gly Thr Gly Lys Ala Trp Asp Trp Glu Val Asp  
 180 185 190  
 Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Val Asp Met  
 195 200 205  
 Asp His Pro Glu Val Ile His Glu Leu Arg Asn Trp Gly Val Trp Tyr  
 210 215 220  
 Thr Asn Thr Leu Asn Ile Asp Gly Phe Arg Ile Asp Ala Val Lys His  
 225 230 235 240  
 Ile Lys Tyr Ser Phe Thr Arg Asp Trp Leu Thr His Val Arg Asn Thr  
 245 250 255  
 Thr Gly Lys Pro Met Phe Ala Val Ala Glu Phe Trp Lys Asn Asp Leu  
 260 265 270  
 Gly Ala Ile Glu Asn Tyr Leu Asn Lys Thr Ser Trp Asn His Ser Val  
 275 280 285  
 Phe Asp Val Pro Leu His Tyr Asn Leu Tyr Asn Ala Ser Asn Ser Gly  
 290 295 300  
 Gly Tyr Tyr Asp Met Arg Asn Ile Leu Asn Gly Ser Val Val Gln Lys  
 305 310 315 320  
 His Pro Thr His Ala Val Thr Phe Val Asp Asn His Asp Ser Gln Pro  
 325 330 335  
 Gly Glu Ala Leu Glu Ser Phe Val Gln Gln Trp Phe Lys Pro Leu Ala  
 340 345 350  
 Tyr Ala Leu Val Leu Thr Arg Glu Gln Gly Tyr Pro Ser Val Phe Tyr  
 355 360 365  
 Gly Asp Tyr Tyr Gly Ile Pro Thr His Gly Val Pro Ala Met Lys Ser  
 370 375 380

Lys Ile Asp Pro Leu Leu Gln Ala Arg Gln Thr Phe Ala Tyr Gly Thr  
 385 390 395 400  
 Gln His Asp Tyr Phe Asp His His Asp Ile Ile Gly Trp Thr Arg Glu  
 405 410 415  
 Gly Asn Ser Ser His Pro Asn Ser Gly Leu Ala Thr Ile Met Ser Asp  
 420 425 430  
 Gly Pro Gly Gly Asn Lys Trp Met Tyr Val Gly Lys Asn Lys Ala Gly  
 435 440 445  
 Gln Val Trp Arg Asp Ile Thr Gly Asn Arg Thr Gly Thr Val Thr Ile  
 450 455 460  
 Asn Ala Asp Gly Trp Gly Asn Phe Ser Val Asn Gly Gly Ser Val Ser  
 465 470 475 480  
 Val Trp Val Lys Gln  
 485

## (2) INFORMATION FOR SEQ ID NO: 2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 485 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

His His Asn Gly Thr Asn Gly Thr Met Met Gln Tyr Phe Glu Trp His  
 1 5 10 15  
 Leu Pro Asn Asp Gly Asn His Trp Asn Arg Leu Arg Asp Asp Ala Ser  
 20 25 30  
 Asn Leu Arg Asn Arg Gly Ile Thr Ala Ile Trp Ile Pro Pro Ala Trp  
 35 40 45  
 Lys Gly Thr Ser Gln Asn Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr  
 50 55 60  
 Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly  
 65 70 75 80  
 Thr Arg Ser Gln Leu Glu Ser Ala Ile His Ala Leu Lys Asn Asn Gly  
 85 90 95  
 Val Gln Val Tyr Gly Asp Val Val Met Asn His Lys Gly Gly Ala Asp  
 100 105 110  
 Ala Thr Glu Asn Val Leu Ala Val Glu Val Asn Pro Asn Asn Arg Asn  
 115 120 125  
 Gln Glu Ile Ser Gly Asp Tyr Thr Ile Glu Ala Trp Thr Lys Phe Asp  
 130 135 140  
 Phe Pro Gly Arg Gly Asn Thr Tyr Ser Asp Phe Lys Trp Arg Trp Tyr  
 145 150 155 160

His Phe Asp Gly Val Asp Trp Asp Gln Ser Arg Gln Phe Gln Asn Arg  
 165 170 175  
 Ile Tyr Lys Phe Arg Gly Asp Gly Lys Ala Trp Asp Trp Glu Val Asp  
 180 185 190  
 Ser Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Val Asp Met  
 195 200 205  
 Asp His Pro Glu Val Val Asn Glu Leu Arg Arg Trp Gly Glu Trp Tyr  
 210 215 220  
 Thr Asn Thr Leu Asn Leu Asp Gly Phe Arg Ile Asp Ala Val Lys His  
 225 230 235 240  
 Ile Lys Tyr Ser Phe Thr Arg Asp Trp Leu Thr His Val Arg Asn Ala  
 245 250 255  
 Thr Gly Lys Glu Met Phe Ala Val Ala Glu Phe Trp Lys Asn Asp Leu  
 260 265 270  
 Gly Ala Leu Glu Asn Tyr Leu Asn Lys Thr Asn Trp Asn His Ser Val  
 275 280 285  
 Phe Asp Val Pro Leu His Tyr Asn Leu Tyr Asn Ala Ser Asn Ser Gly  
 290 295 300  
 Gly Asn Tyr Asp Met Ala Lys Leu Leu Asn Gly Thr Val Val Gln Lys  
 305 310 315 320  
 His Pro Met His Ala Val Thr Phe Val Asp Asn His Asp Ser Gln Pro  
 325 330 335  
 Gly Glu Ser Leu Glu Ser Phe Val Gln Glu Trp Phe Lys Pro Leu Ala  
 340 345 350  
 Tyr Ala Leu Ile Leu Thr Arg Glu Gln Gly Tyr Pro Ser Val Phe Tyr  
 355 360 365  
 Gly Asp Tyr Tyr Gly Ile Pro Thr His Ser Val Pro Ala Met Lys Ala  
 370 375 380  
 Lys Ile Asp Pro Ile Leu Glu Ala Arg Gln Asn Phe Ala Tyr Gly Thr  
 385 390 395 400  
 Gln His Asp Tyr Phe Asp His His Asn Ile Ile Gly Trp Thr Arg Glu  
 405 410 415  
 Gly Asn Thr Thr His Pro Asn Ser Gly Leu Ala Thr Ile Met Ser Asp  
 420 425 430  
 Gly Pro Gly Gly Glu Lys Trp Met Tyr Val Gly Gln Asn Lys Ala Gly  
 435 440 445  
 Gln Val Trp His Asp Ile Thr Gly Asn Lys Pro Gly Thr Val Thr Ile  
 450 455 460  
 Asn Ala Asp Gly Trp Ala Asn Phe Ser Val Asn Gly Gly Ser Val Ser  
 465 470 475 480  
 Ile Asp Val Lys Arg  
 485

## (2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 514 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Ala Ala Pro Phe Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Tyr Leu  
 1 5 10 15  
 Pro Asp Asp Gly Thr Leu Trp Thr Lys Val Ala Asn Glu Ala Asn Asn  
 20 25 30  
 Leu Ser Ser Leu Gly Ile Thr Ala Leu Trp Leu Pro Pro Ala Tyr Lys  
 35 40 45  
 Gly Thr Ser Arg Ser Asp Val Gly Tyr Gly Val Tyr Asp Leu Tyr Asp  
 50 55 60  
 Leu Gly Glu Phe Asn Gln Lys Gly Ala Val Arg Thr Lys Tyr Gly Thr  
 65 70 75 80  
 Lys Ala Gln Tyr Leu Gln Ala Ile Gln Ala Ala His Ala Ala Gly Met  
 85 90 95  
 Gln Val Tyr Ala Asp Val Val Phe Asp His Lys Gly Gly Ala Asp Gly  
 100 105 110  
 Thr Glu Trp Val Asp Ala Val Glu Val Asn Pro Ser Asp Arg Asn Gln  
 115 120 125  
 Glu Ile Ser Gly Thr Tyr Gln Ile Gln Ala Trp Thr Lys Phe Asp Phe  
 130 135 140  
 Pro Gly Arg Gly Asn Thr Tyr Ser Ser Phe Lys Trp Arg Trp Tyr His  
 145 150 155 160  
 Phe Asp Gly Val Asp Trp Asp Glu Ser Arg Lys Leu Ser Arg Ile Tyr  
 165 170 175  
 Lys Phe Arg Gly Ile Gly Lys Ala Trp Asp Trp Glu Val Asp Thr Glu  
 180 185 190  
 Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Leu Asp Met Asp His  
 195 200 205  
 Pro Glu Val Val Thr Glu Leu Lys Ser Trp Gly Lys Trp Tyr Val Asn  
 210 215 220  
 Thr Thr Asn Ile Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys  
 225 230 235 240  
 Phe Ser Phe Phe Pro Asp Trp Leu Ser Asp Val Arg Ser Gln Thr Gly  
 245 250 255  
 Lys Pro Leu Phe Thr Val Gly Glu Tyr Trp Ser Tyr Asp Ile Asn Lys

260	265	270
Leu His Asn Tyr Ile Met Lys Thr Asn Gly Thr Met Ser Leu Phe Asp		
275	280	285
Ala Pro Leu His Asn Lys Phe Tyr Thr Ala Ser Lys Ser Gly Gly Thr		
290	295	300
Phe Asp Met Arg Thr Leu Met Thr Asn Thr Leu Met Lys Asp Gln Pro		
305	310	315
Thr Leu Ala Val Thr Phe Val Asp Asn His Asp Thr Glu Pro Gly Gln		
325	330	335
Ala Leu Gln Ser Trp Val Asp Pro Trp Phe Lys Pro Leu Ala Tyr Ala		
340	345	350
Phe Ile Leu Thr Arg Gln Glu Gly Tyr Pro Cys Val Phe Tyr Gly Asp		
355	360	365
Tyr Tyr Gly Ile Pro Gln Tyr Asn Ile Pro Ser Leu Lys Ser Lys Ile		
370	375	380
Asp Pro Leu Leu Ile Ala Arg Arg Asp Tyr Ala Tyr Gly Thr Gln His		
385	390	395
Asp Tyr Leu Asp His Ser Asp Ile Ile Gly Trp Thr Arg Glu Gly Val		
405	410	415
Thr Glu Lys Pro Gly Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro		
420	425	430
Gly Gly Ser Lys Trp Met Tyr Val Gly Lys Gln His Ala Gly Lys Val		
435	440	445
Phe Tyr Asp Leu Thr Gly Asn Arg Ser Asp Thr Val Thr Ile Asn Ser		
450	455	460
Asp Gly Trp Gly Glu Phe Lys Val Asn Gly Gly Ser Val Ser Val Trp		
465	470	475
480		
Val Pro Arg Lys Thr Thr Val Ser Thr Ile Ala Trp Ser Ile Thr Thr		
485	490	495
Arg Pro Trp Thr Asp Glu Phe Val Arg Trp Thr Glu Pro Arg Leu Val		
500	505	510
Ala Trp		

## (2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1455 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA (genomic)

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

CATCATATG GAACAAATGG TACTATGATG CAATATTCG AATGGTATTG GCCAAATGAC

60

GGGAATCATT GGAACAGGTT GAGGGATGAC GCAGCTAACT TAAAGAGTAA AGGGATAACA	120
GCTGTATGGA TCCCACCTGC ATGGAAGGGG ACTTCCCAGA ATGATGTAGG TTATGGAGCC	180
TATGATTTAT ATGATCTTGG AGAGTTAAC CAGAAGGGGA CGGTTCGTAC AAAATATGGA	240
ACACCCAACCC AGCTACAGGC TGCGGTGACC TCTTAAAAAA ATAACGGCAT TCAGGTATAT	300
GGTGATGTCG TCATGAATCA TAAAGGTGGA GCAGATGGTA CGGAAATTGT AAATGCGGTA	360
GAAGTGAATC GGAGCAACCG AAACCAGGAA ACCTCAGGAG AGTATGCAAT AGAAGCGTGG	420
ACAAAAGTTG ATTTCTTGG AAGAGGAAAT AACCAATTCCA GCTTTAAGTG GCGCTGGTAT	480
CATTTGATG GGACAGATTG GGATCAGTCA CGCCAGCTTC AAAACAAAAT ATATAAATTC	540
AGGGGAACAG GCAAGGCCCTG GGACTGGGAA GTCGATACAG AGAATGGCAA CTATGACTAT	600
CTTATGTATG CAGACGTGGA TATGGATCAC CCAGAAGTAA TACATGAACT TAGAAACTGG	660
GGAGTGTGGT ATACGAATAC ACTGAACCTT GATGGATTAA GAATAGATGC AGTGAACAT	720
ATAAAATATA GCTTTACGAG AGATTGGCTT ACACATGTGC GTAACACCAAC AGGTAAACCA	780
ATGTTGCAG TGGCTGAGTT TTGGAAAAAT GACCTTGGTG CAATTGAAAA CTATTGAAT	840
AAAACAAGTT GGAATCACTC GGTGTTGAT GTTCCTCTCC ACTATAATT GTACAATGCA	900
TCTAATAGCG GTGGTTATTA TGATATGAGA AATATTTAA ATGGTTCTGT GGTGCAAAAA	960
CATCCAAACAC ATGCCGTTAC TTTTGTGAT AACCATGATT CTCAGCCCCG GGAAGCATTG	1020
GAATCCTTGTG TTCAACAATG GTTTAAACCA CTTGCATATG CATTGGTTCT GACAAGGGAA	1080
CAAGGTTATC CTTCCGTATT TTATGGGGAT TACTACGGTA TCCCAACCCA TGGTGTTCG	1140
GCTATGAAAT CTAAAATAGA CCCTCTCTG CAGGCACGTC AARCTTTGC CTATGGTACG	1200
CAGCATGATT ACTTTGATCA TCATGATATT ATCGGTTGGA CAAGAGAGGG AAATAGCTCC	1260
CATCCAAATT CAGGCCTTGC CACCATTATG TCAGATGGTC CAGGTGGTAA CAAATGGATG	1320
TATGTGGGGA AAAATAAAGC GGGACAAGTT TGGAGAGATA TTACCGGAAA TAGGACAGGC	1380
ACCGTCACAA TTAATGCAGA CGGATGGGTT AATTCTCTG TTAATGGAGG GTCCGTTTCG	1440
GTGGGGTGA AGCAA	1455

## (2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1455 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

CATCATAATG GGACAAATGG GACGATGATG CAATACTTGT AATGGCACTT GCCTAATGAT

60

GGGAATCACT GGAATAGATT AAGAGATGAT GCTAGTAATC TAAGAAATAG AGGTATAACC	120
GCTATTTGGA TTCCGCCTGC CTGGAAAGGG ACTTCGAAA ATGATGTGGG GTATGGAGCC	180
TATGATCTTT ATGATTTAGG GGAATTAAAT CAAAAGGGGA CGGTTCGTAC TAAAGTATGGG	240
ACACGTAGTC AATTGGAGTC TGCCATCCAT GCTTAAAGA ATAATGGCGT TCAAQTTTAT	300
GGGGATGTAG TGATGAACCA TAAAGGAGGA GCTGATGCTA CAGAAAACGT TCTTGCTGTC	360
GAGGTGAATC CAAATAACCG GAATCAAGAA ATATCTGGGG ACTACACAAT TGAGGCTTGG	420
ACTAAGTTG ATTTCCAGG GAGGGGTAAAT ACATACTCAG ACTTTAAATG GCCTTGGTAT	480
CATTCGATG GTGTAGATTG GGATCAATCA CGACAATTCC AAAATCGTAT CTACAAATTC	540
CGAGGTGATG GTAAGGCATG GGATTGGAA GTAGATTGG AAAATGGAAA TTATGATTAT	600
TTAATGTATG CAGATGTAGA TATGGATCAT CCGGAGGTAG TAAATGAGCT TAGAAGATGG	660
GGAGAATGGT ATACAAATAC ATTAAATCTT GATGGATTAA GGATCGATGC GGTGAAGCAT	720
ATTAAATATA GCTTTACACG TGATTGGTTG ACCCATGTAA GAAACGCAAC GGGAAAAGAA	780
ATGTTGCTG TTGCTGAATT TTGGAAAAAT GATTTAGGTG CCTTGGAGAA CTATTAAAT	840
AAAACAAACT GGAATCATTG TGTCTTGAT GTCCCCCTTC ATTATAATCT TTATAACGCG	900
TCAAATAGTG GAGGCAACTA TGACATGGCA AAACCTCTTA ATGGAACGGT TGTTCAAAAG	960
CATCCAATGC ATGCCGTAAAC TTTTGTGGAT AATCACGATT CTCAACCTGG GGAATCATTAA	1020
GAATCATTG TACAAGAATG GTTTAAGCCA CTTGCTTATG CGCTTATTAA ACAAGAGAA	1080
CAAGGCTATC CCTCTGTCTT CTATGGTAC TACTATGGAA TTCCAACACA TAGTGTCCC	1140
GCAATGAAAG CCAAGATTGA TCCAATCTTA GAGGCACGTC AAAATTTGC ATATGGAACA	1200
CAACATGATT ATTTGACCA TCATAATATA ATCGGATGGA CACGTGAAGG AAATACCACG	1260
CATCCCAATT CAGGACTTGC GACTATCATG TCGGATGGC CAGGGGGAGA GAAATGGATG	1320
TACGTAGGGC AAAATAAAGC AGGTCAAGTT TGGCATGACA TAACTGGAAA TAAACCAGGA	1380
ACAGTTACGA TCAATGCAGA TGGATGGGCT AATTTCAG TAAATGGAGG ATCTGTTCC	1440
ATTTGGGTGA AACGA	1455

## (2) INFORMATION FOR SEQ ID NO: 6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1548 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GCCGCACCGT TTAACGGCAC CATGATGCAG TATTTGAAT GGTACTTGCC GGATGATGGC

60

ACGTTATGGA CCAAAGTGGC CAATGAAGCC AACAACTTAT CCAGCCTTGG CATCACCGCT	120
CTTTGGCTGC CGCCCCCTTA CAAAGGAACA AGCCGCAGCG ACGTAGGGTA CGGAGTATAAC	180
GACTTGTATG ACCTCGGCGA ATTCAATCAA AAAGGGACCG TCCGCACAAA ATACGGAACA	240
AAAGCTCAAT ATCTTCAAGC CATTCAAGCC GCCCACGCCG CTGGATGCA AGTGTACGCC	300
GATGTCGTGT TCGACCATAA AGGCGGCGCT GACGGCACGG AATGGGTGGA CGCCGTCGAA	360
GTCAATCCGT CCGACCGCAA CCAAGAAATC TCGGGCACCT ATCAAATCCA AGCATGGACG	420
AAATTGATT TTCCCGGGCG GGGCAACACC TACTCCAGCT TTAAGTGGCG CTGGTACCAT	480
TTTGACGGCG TTGATTGGGA CGAAAGCCGA AAATTGAGCC GCATTTACAA ATTCCGGCGC	540
ATCGGCAAAG CGTGGGATTG GGAAGTAGAC ACGGAAAACG GAAACTATGA CTACTTAATG	600
TATGCCGACC TTGATATGGA TCATCCGAA GTCGTACCG AGCTGAAAAA CTGGGGAAA	660
TGGTATGTCA ACACAACGAA CATTGATGGG TTCCGGCTTG ATGCCGTCAA GCATATTAAG	720
TTCAGTTTT TTCTCGATTG GTTGTGATG GTGCCTTCTC AGACTGGCAA GCCGCTATTT	780
ACCGTCGGGG AATATTGGAG CTATGACATC AACAAAGTTGC ACAATTACAT TACGAAAACA	840
GACGGAACGA TGTCTTGTT TGATGCCCG TACACACAACA AATTTTATAC CGCTTCCAAA	900
TCAGGGGGCG CATTGATAT GCGCACGTTA ATGACCAATA CTCTCATGAA AGATCAACCG	960
ACATTGGCCG TCACCTTCGT TGATAATCAT GACACCGAAC CCGGCCAAGC GCTGCAGTCA	1020
TGGGTCGACC CATGGTCAA ACCGTTGGCT TACGCCCTTA TTCTAACTCG GCAGGAAGGA	1080
TACCCGTGCG TCTTTATGG TGACTATTAT GGCATTCCAC AATATAACAT TCCTCGCTG	1140
AAAAGCAAA TCGATCCGCT CCTCATCGCG CGCAGGGATT ATGCTTACGG AACGCAACAT	1200
GATTATCTTG ATCACTCCGA CATCATCGGG TGGACAAGGG AAGGGGGCAC TGAAAAACCA	1260
GGATCCGGAC TGGCCGCACT GATCACCGAT GGGCCGGGAG GAAGCAAATG GATGTACGTT	1320
GGCAAACAAAC ACGCTGGAAA AGTGTCTAT GACCTTACCG GCAACCGGAG TGACACCGTC	1380
ACCATCAACA GTGATGGATG GGGGAATTC AAAGTCAATG GCGGTTCGGT TTCGGTTGG	1440
GTTCCCTAGAA AAACGACCGT TTCTACCATC GCTCGGCCGA TCACAACCCG ACCGTGGACT	1500
GGTGAATTGCG TCCGTTGGAC CGAACACCGG TTGGTGGCAT GCCCTGAA	1548

## (2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 485 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

His His Asn Gly Thr Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Tyr  
 1 5 10 15  
 Leu Pro Asn Asp Gly Asn His Trp Asn Arg Leu Asn Ser Asp Ala Ser  
 20 25 30  
 Asn Leu Lys Ser Lys Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Trp  
 35 40 45  
 Lys Gly Ala Ser Gln Asn Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr  
 50 55 60  
 Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly  
 65 70 75 80  
 Thr Arg Ser Gln Leu Gln Ala Ala Val Thr Ser Leu Lys Asn Asn Gly  
 85 90 95  
 Ile Gln Val Tyr Gly Asp Val Val Met Asn His Lys Gly Ala Asp  
 100 105 110  
 Ala Thr Glu Met Val Arg Ala Val Glu Val Asn Pro Asn Asn Arg Asn  
 115 120 125  
 Gln Glu Val Thr Gly Glu Tyr Thr Ile Glu Ala Trp Thr Arg Phe Asp  
 130 135 140  
 Phe Pro Gly Arg Gly Asn Thr His Ser Ser Phe Lys Trp Arg Trp Tyr  
 145 150 155 160  
 His Phe Asp Gly Val Asp Trp Asp Gln Ser Arg Arg Leu Asn Asn Arg  
 165 170 175  
 Ile Tyr Lys Phe Arg Gly His Gly Lys Ala Trp Asp Trp Glu Val Asp  
 180 185 190  
 Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Ile Asp Met  
 195 200 205  
 Asp His Pro Glu Val Val Asn Glu Leu Arg Asn Trp Gly Val Trp Tyr  
 210 215 220  
 Thr Asn Thr Leu Gly Leu Asp Gly Phe Arg Ile Asp Ala Val Lys His  
 225 230 235 240  
 Ile Lys Tyr Ser Phe Thr Arg Asp Trp Ile Asn His Val Arg Ser Ala  
 245 250 255  
 Thr Gly Lys Asn Met Phe Ala Val Ala Glu Phe Trp Lys Asn Asp Leu  
 260 265 270  
 Gly Ala Ile Glu Asn Tyr Leu Gln Lys Thr Asn Trp Asn His Ser Val  
 275 280 285  
 Phe Asp Val Pro Leu His Tyr Asn Leu Tyr Asn Ala Ser Lys Ser Gly  
 290 295 300  
 Gly Asn Tyr Asp Met Arg Asn Ile Phe Asn Gly Thr Val Val Gln Arg  
 305 310 315 320  
 His Pro Ser His Ala Val Thr Phe Val Asp Asn His Asp Ser Gln Pro  
 325 330 335

Glu Glu Ala Leu Glu Ser Phe Val Glu Glu Trp Phe Lys Pro Leu Ala  
 340 345 350  
 Tyr Ala Leu Thr Leu Thr Arg Glu Gln Gly Tyr Pro Ser Val Phe Tyr  
 355 360 365  
 Gly Asp Tyr Tyr Gly Ile Pro Thr His Gly Val Pro Ala Met Arg Ser  
 370 375 380  
 Lys Ile Asp Pro Ile Leu Glu Ala Arg Gln Lys Tyr Ala Tyr Gly Lys  
 385 390 395 400  
 Gln Asn Asp Tyr Leu Asp His His Asn Ile Ile Gly Trp Thr Arg Glu  
 405 410 415  
 Gly Asn Thr Ala His Pro Asn Ser Gly Leu Ala Thr Ile Met Ser Asp  
 420 425 430  
 Gly Ala Gly Gly Ser Lys Trp Met Phe Val Gly Arg Asn Lys Ala Gly  
 435 440 445  
 Gln Val Trp Ser Asp Ile Thr Gly Asn Arg Thr Gly Thr Val Thr Ile  
 450 455 460  
 Asn Ala Asp Gly Trp Gly Asn Phe Ser Val Asn Gly Gly Ser Val Ser  
 465 470 475 480  
 Ile Trp Val Asn Lys  
 485

## (2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 30 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GCTGCGGTGA CCTCTTTAAA AAATAACGGC

30

## (2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CCACCGCTAT TAGATGCATT GTAC

24

## (2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 32 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CTTACGTATG CAGACGTCGA TATGGATCAC CC

32

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GATCCATATC GACGTCTGCA TACGTAAGAT AGTC

34

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TTASGGGCAA GGCCTGGGAC TGG

23

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CCCAGGGCTT GCCCSTAAAT TTATATATTT TGTTTG

37

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GGTTTCGGTT CGAAGGATTG ACTTCTACCG C

31

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GCGGTAGAAG TGAATCCTTC GAACCGAAAC CAG

33

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 43 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GGTACTATCG TAACAATGGC CGATTGCTGA CGCTGTTATT TGC

43

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 28 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CTGTGACTGG TGAGTACTCA ACCAAGTC

28

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 35 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CTACTTCCCA ATCCCAGCT TTACCTCGGA ATTG

35

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 35 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CAAATTCCGA GGTAAAGCTT GGGATTGGGA AGTAG

35

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 24 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TTGAACAAACC GTTCCATTAA GAAG

24

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 60 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CTCTGTATCG ACTTCCCAGT CCCAAGCTTT TGTCCTGAAT TTATATATTT TGTTTGAG

60

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 60 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CTCTGTATCG ACTTCCCAGT CCCAAGCTTT GCCTCCGAAT TTATATATTT TGTTTGAG

60

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 51 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

ATGTGTAAGC CAATCGCGAG TAAAGCTAAA TTTTATATGT TTCACTGCAT C

51

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 34 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GCACCAAGGT CATTTCGCCA GAATTCAAGCC ACTG

34

## (2) INFORMATION FOR SEQ ID NO:25:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

TGTCAAGAAC AACGCGTATG CACATGGTTT AAACCATTG

39

## (2) INFORMATION FOR SEQ ID NO:26:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

ACCACCTGGA CCATCGCTGC AGATGGTGGC AAGGCCTGAA TT

42

## (2) INFORMATION FOR SEQ ID NO:27:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GGCAAAGTT TGACGTGCCT CGAGAAGAGG GTCTAT

36

## (2) INFORMATION FOR SEQ ID NO:28:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

TTGTCCCGCT TTATTCTGGC CAACATACAT CCATT

36

## (2) INFORMATION FOR SEQ ID NO:29:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CCCAATCCCC AGCTTTACCA YCGAACTTGT AGATAACG

37

## (2) INFORMATION FOR SEQ ID NO:30:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CCCAATCCA AGCTTTATCT CSGAACTTGT AGATACG

37

## (2) INFORMATION FOR SEQ ID NO:31:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GATCCATATC GACGTCTGCA TACAGTAAAT AATC

34

## (2) INFORMATION FOR SEQ ID NO:32:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

GATCCATATC GACGTCTGCA TAAATTAAAT AATC

34